Journal of Applied Economics, Vol. VII, No. I (May 2004), 163-193

STATE EXECUTIONS, DETERRENCE, AND THE INCIDENCE OF MURDER

PAUL R. ZIMMERMAN*
Federal Communications Commission

Submitted April 2003; accepted February 2004

This study employs a panel of U.S. state-level data over the years 1978-1997 to estimate the deterrent effect of capital punishment. Particular attention is paid to problems of endogeneity bias arising from the non-random assignment of death penalty laws across states and a simultaneous relationship between murders and the deterrence probabilities. The primary innovation of the analysis lies in the estimation of a simultaneous equations system whose identification is based upon the employment of instrumental variables motivated by the theory of public choice. The estimation results suggest that structural estimates of the deterrent effect of capital punishment are likely to be downward biased due to the influence of simultaneity. Correcting for simultaneity, the estimates imply that a state execution deters approximately fourteen murders per year on average. Finally, the results also suggest that the announcement effect of capital punishment, as opposed to the existence of a death penalty provision, is the mechanism actually driving the deterrent effect associated with state executions.

JEL classification codes: K42, H00
Key words: capital punishment, deterrence, executions, murder

I. Introduction

In January 2000 then-Illinois Governor George Ryan commuted the death

---

* Senior Economist, Federal Communications Commission, 445 12th Street SW, Washington, DC, 20554. The author would like to thank an anonymous referee for many useful comments and suggestions. The views expressed in this paper are those of the author and not necessarily those of the Federal Communications Commission, its Chairman or Commissioners, or any other staff. All errors are my own. E-mail: paul.r.zimmerman@att.net.

sentences of all 167 inmates on death row in the state to life imprisonment.[1] Ryan, who made his decision just days before his term was set to expire, based his decision on apparent flaws in the capital punishment system which led to several convicted offenders being taken off death row due to ambiguities regarding their guilt. State prosecutors, victims' families, and even some politicians decried the Governor's decision in response, arguing that the decision was capricious and might even lead to an increase in murders as potential offenders would no longer fear the possibility of receiving a death sentence.

The question of whether executions can serve to deter capital murders has been given extensive attention by numerous academic disciplines. Research by economists into this question originates with the seminal work of Ehrlich (1975, 1977). By extending Becker's (1968) theory of the rational offender Ehrlich develops a positive approach towards testing the deterrence hypothesis using multiple regression techniques. His results, based upon both national time-series and state cross-sectional data, found a large and statistically significant deterrent effect of executions.

Ehrlich's findings generated a legion of subsequent empirical studies on the deterrent effect of capital punishment that questioned the validity of his results based upon data quality and econometric specification issues (among others).[2] Indeed, one of the primary challenges posed to Ehrlich's work was his use of simultaneous equations procedures and the instruments employed in identifying his empirical model.[3] Consequently, for nearly every study that apparently refuted Ehrlich's results there arose another that seemed to support them.[4] As such, to this day no definitive consensus among economists (and many other social scientists) on the deterrent effect of capital punishment exists. As explained below, since the effect of state executions on murder is

---

[1] In addition to the commutations, a moratorium was placed on all executions pending a review of the state's capital punishment system.

[2] See Cameron (1994) for a comprehensive review and critique of these studies.

[3] See, e.g., Blumstein et al. (1978) and Brier and Feinberg (1980).

[4] For recent support of the deterrence hypothesis see Brumm and Cloninger (1996), Dezhbakhsh et al. (2003), and Mocan and Gittings (2003). Recent studies which refute the deterrence hypothesis include Grogger (1990), Andreoni (1995), and Katz et al. (2003).

theoretically ambiguous, the question of whether or not capital punishment actually deters murder ultimately depends upon empirical analysis. Indeed, an empirical determination of the deterrent effect of capital punishment is important given that most credible estimates of the value of life fall into the $3 million to $7 million dollar range [Viscusi (1993)]. As such, the potential social benefits that could be realized from even a small deterrent effect of capital punishment might be substantial.

This paper estimates the deterrent effect of capital punishment using state-level panel data from the post-moratorium period (1977-1997). Panel data (i.e., pooled cross-section and time series data) has several distinct advantages relative to ordinary time-series or cross-sectional data. For instance, panel data provides more degrees of freedom and allows for more robust estimation relative to pure time-series or cross-sectional approaches. More importantly, certain immeasurable (i.e., omitted) factors may determine the rate of murders and the rate of executions simultaneously.[5] However, panel data allows for the estimation of fixed-effects models that control for the influences of unobserved state or year-specific heterogeneity.

Another issue that confounds estimation of the deterrent effect of capital punishment is a possible simultaneous relationship between the relevant deterrence probabilities and the rate of murder.[6] The primary innovation of

[5] As such, a spurious negative correlation between the rates of execution and murder might arise, for instance, if states that have relatively strong underlying anti-crime sentiments (leading to few murders) are also states that tend to conduct executions. Such effects in turn might lead one to conclude that a deterrent effect of capital punishment exists even though it does not.

[6] The term deterrence probabilities refers collectively to the probability of being arrested for murder, the probability of being sentenced to death conditional upon arrest for murder, and the probability of being executed conditional on being sentenced to death throughout. The use of these three measures might not be inclusive of all the relevant probabilities underlying the imposition of an execution. For instance, before an arrest can be made the crime must be either reported to the police or discovered (although in the case of murder most offenses are reported). In addition, the probability of being sentenced to death given an arrest is actually determined by a series of probabilities including the probability of being prosecuted given an arrest (note that some arrests might not be prosecuted due to mistakes in arrests, the arrest of persons who are later determined to be innocent, etc.), the probability of conviction given prosecution, and the probability of a death sentence given conviction. Data limitations preclude the consideration of all the potentially relevant

this paper is in estimating the deterrent effect of state executions through the use of a simultaneous equations model [based upon the original theory of Ehrlich (1975)] which employs a set of instrumental variables derived from the theory of public choice as it relates to the criminal justice system and bureaucratic behavior. The instruments include the proportion of annual state murders committed by strangers, the lagged number of prisoners released from death row, and the number of state executions that were botched in the previous year, among several others. The use of these particular instruments has not been considered in the previous literature, and it is shown that they provide plausibly exogenous and significant sources of variation in the key endogenous variables of the model by satisfying the two main criteria of instrument selection (i.e., relevancy and validity).

The remainder of the paper proceeds as follows. Section II discusses the data and theoretical rationale for employing a simultaneous equations estimation procedure to determine the deterrent effect of capital punishment. In addition, this section discusses the motivation for employing instrumental variables based upon the theory of public choice as pertaining to the criminal justice system and bureaucratic behavior to identify the unbiased effect of state executions on the rate of murder. Section III presents the results of the empirical estimation. It is shown that ordinary least squares estimates of the deterrent effect of capital punishment are likely to suffer from simultaneity bias (i.e., the relevant deterrence measures are likely endogenous). This result implies that the structural estimates of the deterrent effect of state executions are downward biased. Correcting for the influence of simultaneity bias, it is estimated that each state execution deters approximately fourteen murders per year on average. Section IV discuses implications arising from the differences between instrumented and purely structural estimates of the deterrent effect of capital punishment and discusses some implications for further research. Section V concludes.

---

conditional probabilities. However, the measures considered here are similar to those employed in other studies [e.g., Ehrlich (1975), Dezhbakhsh et al. (2003)].

II. Empirical Specification and Data

A panel of U.S. state-level data is employed to examine the relationship between the deterrence probabilities and state per-capita murder rates. The panel covers the years 1978 to 1997 for the 50 states (excluding Washington, D.C.). The beginning and ending dates of the data set were selected for three reasons. First, there were no executions in the United States between the years 1972 and 1976 due to the moratorium on state executions established by the Supreme Court in *Furman vs. Georgia*, 1972. Second, state arrest data are not available prior to 1977. Third, several of the variables are derived from data published by the Bureau of Justice Statistics and are not available for more recent years.

Similarly, several factors motivated the use of data aggregated to the state-level as opposed to data aggregated to the county level to estimate the deterrent effect of state executions. First, and most fundamentally, capital punishment laws are only enacted by state statute and death sentences are only handed down by courts at the state-level of jurisdiction. As such, there is no real cross-county variation in the number of persons sentenced to death or executed in a given year. Second, while Lott and Mustard (1997) advocate the use of county-level crime data since most of the variation in crimes occurs at the county level as opposed to state level, counts of murder and murder arrests in county data contain a large number of zero-valued observations. Indeed, Plassmann and Tideman (2001) note that dropping all counties from Lott and Mustard's (1997) data set with zero reported murders leads to a loss of over forty percent of their sample size. Plassmann and Tideman (2001, p. 774) also argue: "The very large proportions of zeros in the cases of murders, rapes, and robberies imply that valid distributions of these data sets will have substantial mass points at zero for many counties, and an adequate statistical analysis of these data must be taken into account." As such, the use of classical regression techniques applied to county crime data is likely to be inappropriate in estimating the deterrent effect of state executions. On the other hand, state-level data do not suffer from the problem of large numbers of murder or murder arrest data containing zero recorded observations in any year. Finally, there are simply fewer data problems associated with state-level data.[7] For example, reported crime and arrest records at the county level are often incomplete and those collected by the National Archive of Criminal Justice Data changed procedures for correcting missing data over the sample period used in this paper [e.g., see Duggan (2001) and Marvell and Moody (2001)].[8]

It is assumed that state per-capita murder rates are determined according to the following structural model:

$$\textit{Murders Per Capita}_{i,t} = \beta_0 + \beta_1 \Pr(\mathrm{a})_{i,t} + \beta_2 \Pr(\mathrm{c} \mid \mathrm{a})_{i,t} + \beta_3 \Pr(\mathrm{e} \mid \mathrm{c})_{i,t} \qquad (1)$$

$$+ \Phi\mathbf{E}_{i,t} + \Gamma\mathbf{X}_{i,t} + \Omega\mathbf{C}_{i,t} + \theta_i + \lambda_t + \tau_{i,t} + \varepsilon_{i\,t}$$

where the subscript i corresponds to states and t indexes years. The dependent

[7] Duggan (2001) notes that county-level crime data are substantially noisier than state-level data, with the former having standard deviations in the range two to three times those of the latter.

[8] On the other hand, samples of urban areas or multiple county areas such as judicial districts might be more appropriate for crime studies since they are more complete than individual county data and less susceptible to the endogeneity problems associated with state-level data. However, the use of fixed-effects and simultaneous equations models in this study serves to alleviate the concerns associated with relying on state-level observations. In addition, Maltz and Targonski (2002, 2003) note that state-level crime data are not affected to the same extent as county-level crime data by missing observations. This is because often the largest reporting agency in a given county has missing data, which in turn can have a large impact on the county's reported crime rate. However, the largest reporting agencies in a given state rarely have missing data. In addition, state-level crime data files provided by the FBI and BJS take missing data into account by imputing all missing agency data while the NACJD county-level files only impute missing data only if an agency provides at least six months of data. While these missing data problems might not be of a particular concern for reported murders, the models estimated in this study also include measures of crime categories which are expected to be determinants of murders and far more susceptible to the problems of missing observations in county-level data. In addition, including the arrest rate as an independent variable using county-level data forces the researcher to drop many observations from the sample. This is because the arrest rate is defined as the ratio of arrests to reported offenses, and since many counties will have no murders in a given year this measure will be undefined (due to the zero denominator). This truncation of the data can potentially bias the results of the regression estimation [Donohue (2003)].

variable is measured as the number of reported annual murders by state (taken from the FBI's *Uniform Crime Reports*) per 100,000 state residents. β_0 denotes the regression constant and ε the randomly distributed error term. The variables Pr(a), Pr(c|a), and Pr(e|c) are the relevant deterrence probabilities defined as the probability of arrest for murder, the probability of being sentenced to death given arrest for murder, and the probability of being executed given receipt of a death sentence, respectively. The variable **E** is a vector of law-enforcement covariates that includes the per-capita police employment and per-capita prison populations. The variable **X** is a vector of economic and demographic covariates typically included in state-level studies of crime. Economic covariates include the state unemployment rate, poverty rate, and level of per-capita income. State demographic covariates are the percentage of the population that is black, the percentage residing in metropolitan areas, and age structure variables for the percentage of the population ages 18-24, 25-44, 45-64, and over 65 years. Given that some homicides arise during the commission of other crimes we include the vector **C** which contains the per-capita rates of aggravated assault and robbery.

The variables θ and λ denote vectors of state and year dummies respectively. These dummies control for unobserved heterogeneity associated with time-invariant factors that vary across states (i.e., that are state-specific) and time-variant year-specific factors that effect all states symmetrically (e.g., inflation, exposure to violent media, etc.) respectively. The panel aspects of the data are further exploited by inclusion of τ, a vector of state-specific linear time trends. Inclusion of state trend variables eliminates unobservable variation in murder rates within-state that are caused by state-specific factors changing over time. The inclusion of both state and year dummies and state-specific time trends is likely to eliminate the influence of a large number of unobserved factors [Raphael and Winter-Ebmer (2001)]. Equation (1) is estimated using a linear functional form since various observations of the deterrence probabilities take zero values (thereby precluding the use of logarithmic transforms of these measures).[9] Finally, Equation (1) is estimated via ordinary

[9] Previous studies which, for instance, use double-logarithmic specifications in estimating the per-capita murder equation are forced to deal with the issue of zero-valued observations on the deterrence probabilities through various ad hoc procedures, such as assigning them an arbitrarily small positive value or a value of one.

least squares (OLS) where state populations are used in constructing the weights. The use of these weights serves to mitigate the influence of heteroskedasticity caused by greater per-capita murder variation in the smaller states.

Two different constructs of each of the three deterrence probabilities are employed in the estimation. It must be emphasized that the constructed deterrence probabilities employed in this analysis are to be regarded as proxies for the subjective (i.e., perceived) probabilities calculated by potential offenders in deciding whether to commit murder. As such, one must make assumptions on how potential offenders are likely to calculate their subjective probabilities (i.e., process the relevant information pertaining to their probability of being apprehended and subsequently executed for committing murder). Hugo Bedau (1997), a leading opponent of the death penalty, finds an average lag of approximately six years between the time convicted offenders are sentenced to death and the time they are actually executed. As such, on possible way to measure Pr(e|c) would be to divide the current years' executions by the number of death sentences handed down by the courts six years earlier. However, there are several problems associated with constructing Pr(e|c) in this manner. First, using such a measure would necessarily result in a large number of observations from being dropped from the sample. Second, the average time between sentencing and execution across all states may not provide an accurate estimate of the subjective probability potential murderers are most likely to respond to, namely the lag relevant to the state in which they commit their offense.[10]

In addition, Sah (1991) argues that potential criminals are likely to base their subjective probability assessments of punishments upon their interactions with their closest peers. Along a similar line of reasoning, this paper posits that any truly meaningful (subjective) assessment a potential murderer makes on any of the three deterrence probabilities is likely to be based upon the most recent information available to him/her. That is, in making their subjective

[10] Note that the actual application of the death penalty various widely across the executing states in the sample. Indeed, Connecticut, Kansas, New Hampshire, New Jersey, and New York all have death penalty laws but have not executed a singe individual over the sample period. However, each of these states routinely sentences convicted offenders to death.

calculations of the three deterrence probabilities potential murderers are more likely to be aware of judicial action that occurred in the current or previous year than they are of judicial action taken multiple years prior (i.e., assuming any degree of awareness at all). As such, this paper constructs all three deterrence probabilities based upon this fact. Since all observations in the data set correspond to annual values by state, the deterrence probabilities are constructed as either a mix of contemporaneous and once-lagged annual values (referred to hereafter as contemporaneous probabilities) or entirely in terms of once-lagged annual values (referred to hereafter as lagged probabilities) of the variables used to construct the three deterrence probabilities.

Table 1 summarizes the construction of the proxies for the subjective deterrence probabilities employed in this paper. Following most previous econometric studies on the economics of crime, the contemporaneous measure of Pr(a) is defined exclusively in terms of current-year values. However, in the contemporaneous measures of Pr(c|a) and Pr(e|c) it is assumed that individuals associate current-year death sentences and executions with the last complete year for which arrests and sentences (the respective denominators) are known (i.e., the once-lagged values).

Due to the fact that some death penalty states issued no death sentences in particular years or had missing arrest data, some values of Pr(c|a) and Pr(e|c) are undefined (i.e., have missing values for the respective denominators). A backward-looking correction was employed to correct for these undefined observations. Specifically, the most recent year's value for which the probability was defined by the state is used.[11] This method of correction results in an additional 189 observations in the sample used to estimate Equation (1) employing the contemporaneous probabilities and an additional 198 observations in the sample employing the lagged probabilities. The deterrence probabilities which incorporate this correction are referred to as the adjusted probabilities and those that do not as the unadjusted probabilities throughout.

A. The Interaction between Murder and the Contemporaneous Deterrence Probabilities

One implication of Ehrlich's (1975) model of optimal law enforcement activity is that an increase in the rate of murder will induce an increase in the

Table 1. Constructed Deterrence Probabilities

Deterrence probability	Contemporaneous model	Lagged model
Probability of arrest [Pr(a)]	$\frac{\textit{Murder Arrests}_{i,t}}{\textit{Murders}_{i,t}}$	$\frac{\textit{Murder Arrests}_{i,t-1}}{\textit{Murders}_{i,t-1}}$
Probability of conviction given arrest [Pr(c|a)]	$\frac{\textit{Death sentences}_{i,t}}{\textit{Murder arrests}_{i,t}}$	$\frac{\textit{Death sentences}_{i,t-1}}{\textit{Murder arrests}_{i,t-1}}$
Probability of execution given conviction [Pr(e|c)]	$\frac{\textit{Executions}_{i,t}}{\textit{Death sentences}_{i,t}}$	$\frac{\textit{Executions}_{i,t-1}}{\textit{Death sentences}_{i,t-1}}$

Note: *i* indexes states and *t* years. *t-1* denotes the once-lagged (annual) value.

optimal values of the deterrence probabilities, i.e., a simultaneous relationship will exist between the two.[12] As such, the effect of simultaneity in this case is to cause the estimated deterrent effect of capital punishment to be biased downwards (since reverse causation operates in the positive direction). Correcting for this simultaneity bias would therefore be expected to result in more negative estimates of the deterrent effect of capital punishment (i.e., a stronger deterrent effect).

It could also be argued that in states with the death penalty, offenders might commit more murders for purposes of evading the ultimate punishment.

[11] For example, the contemporaneous value of Pr(e|c) for Delaware in 1995 would be calculated as the ratio of executions administered in Delaware in 1995 to the number of death sentences handed down in the state in 1994. However, Delaware did not sentence any offenders to death in 1994, which implies a denominator of zero in the calculation of Pr(e|c). However, the contemporaneous value Pr(e|c) is defined for Delaware in 1994, and is calculated as the number of executions administered in 1994 divided by the number of death sentences handed down in 1993 [= (1)/(6) = 0.1667]. As such, a value of 0.1667 is substituted for the 1995 contemporaneous value of Pr(e|c) in Delaware.

[12] Note that since murders constitute only a small fraction of total crimes it is unlikely that simultaneity bias is an issue with respect to the inclusion of police employment and prison populations in the structural crime equation.

For instance, the probability that an offender is executed is dependent on the probability that he/she is apprehended for their offense. The probability of apprehension in turn is partially dependent on witnesses or other individuals who have information regarding the offender's crime. As such, the response of a rational offender might be to eliminate potential witnesses who could otherwise provide incriminating evidence against him/her and thus increase his/her probability of being executed.[13] If such a lethality effect of capital punishment is operative, estimates of the deterrent effect of capital punishment might be biased upwards since reverse causation operates in the negative direction. Correcting for simultaneity in this case would result in a smaller estimated deterrent effect. As such, note that the theoretical effect of state executions on the incidence of murder cannot be determined a priori. Therefore, one must rely upon empirical methods.

To purge the potential correlation between the structural error term in Equation (1) and the contemporaneous deterrence probabilities requires the use of instrumental variables within the context of a simultaneous-equations estimation approach to identify the uncontaminated (i.e., free from simultaneity bias) effect of the respective deterrence probabilities on the per-capita incidence of murder. Given the assumed linear structure of the model and the primary interest in estimation of Equation (1), the method of two-stage least squares (2SLS) is employed. As such, each contemporaneous deterrence probability in Equation (1) requires its own first-stage equation and at least one unique instrument that is excluded from the structural murder equation that allows for its identification. Since there are three potentially endogenous variables in Equation (1), estimation of a simultaneous equations system consisting of Equation (1) and three separate first-stage regressions requires at least three different instruments. However, the use of an expanded instrument set has several important advantages. First, one can typically capture a greater proportion of the variation in the endogenous regressors by employing more than the minimum number of identifying instruments. This will tend to increase the precision of the coefficient estimates. Second, having more instruments

---

[13] See Marvell and Moody (2001) for evidence of a lethality effect with respect to the impact of "three-strikes" sentencing laws on the incidence of state murders.

than endogenous variables allows for a joint test of model specification and the exogeneity of the constructed instrument set.[14]

B. Public Choice Theory and the Selection of Instruments

The selection of instrumental variables employed in this paper is based upon the theory of public choice as it pertains to the criminal justice system and bureaucratic behavior. However, these arguments are not considered to be so sufficiently compelling that they warrant ignoring the appropriate testing of the instruments excluded from the structural per-capita murder equation. The simultaneous equations system consists of Equation (1) and the following first-stage equations:

$$\Pr(a)_{i,t} = \gamma_0 + \gamma_1 \mathit{Murders\ Per\ Capita}_{i,t} + \gamma_2 (\%\ \mathit{Strangers})_{i,t} + \gamma_3 (\%\ \mathit{Strangers})_{i,t-1} + \theta_i + \lambda_t + \tau_{i,t} + \upsilon_{i,t} \quad (2)$$

$$\Pr(c \mid a)_{i,t} = \delta_0 + \delta_1 \mathit{Murders\ Per\ Capita}_{i,t} + \delta_2 (\%\ \mathit{Non\ Felony})_{i,t} + \delta_3 (\%\ \mathit{Non\ Felony})_{i,t-1} + \delta_4 (\%\ \mathit{Non\ White})_{i,t} + \delta_5 (\%\ \mathit{Non\ White})_{i,t-1} + \theta_i + \lambda_t + \tau_{i,t} + \eta_{i,t} \quad (3)$$

[14] Obtaining unbiased estimates from a 2SLS estimation method requires two conditions on the set of identifying instruments. First, the instruments must be valid, i.e., the instruments must be orthogonal to the structural error term in equation (1). This means that that the instruments only effect per-capita murders indirectly (i.e., through their influence on the endogenous regressors) and can therefore be justifiably excluded from the structural murder regression. Second, the instruments must be relevant. That is, the instruments must have sufficient explanatory power to account for the variation in the (potentially) endogenous execution probabilities. While it is well known that the bias in 2SLS estimates approaches that of OLS as the explanatory power between the instruments and the endogenous regressors approaches zero, recent research has highlighted the particular importance of instrument relevance in finite samples. In particular, Bound et al. (1995) show that 2SLS estimates can be highly biased in small samples if the instruments are weak and have even a small correlation with the structural error term. As such, both the validity of the instruments and their relevancy must be regarded in estimating the deterrent effect of capital punishment through a simultaneous equations method.

$$\Pr(e \mid c) = \varphi_0 + \varphi_1 \, \textit{Murders Per Capita}_{i,t} + \varphi_2 \, \textit{Release}_{i,t\text{-}1} + \varphi_3 \, \textit{Botch}_{i,t\text{-}1} + \theta_i + \lambda_t + \tau_{i,t} + \omega_{i,t} \quad (4)$$

where υ, η, and ω denote randomly-distributed error terms. The instruments used to identify Equations (2) and (3) are taken from the FBI's *Supplementary Homicide Report* (SHR). The SHR is an annual supplement to the FBI's *Uniform Crime Reports* that provides incident-level details on the location, victim, and offender characteristics of reported homicides during the year. These incident-level data are aggregated to the state-level to construct the various instruments. The coverage of the SHR is relatively high; approximately 92% of all homicides over the sample period are detailed. To account for undetailed murders the individual state-year observations are weighted to match the UCR estimated annual state homicide counts.[15]

The instrument used to identify the probability of arrest for murder, Equation (2), is the percentage of state murders that are committed by strangers. These are defined as murders where the offender is not identified as a family member, acquaintance, or as having an unknown relationship to the victim. There are several reasons why this instrument is likely to influence the probability of arrest. First, police are simply less likely to be able to identify an offender that has no relation to the victim and thus less likely to make an arrest. Second, to the extent that crimes committed by strangers are more difficult (and thus more costly) for police to solve, the incentive of police bureaucrats might be to direct resources to easier crimes. For instance, one common measure of police performance employed in the budget negotiation process is the number of arrests made in a given year [Benson et al. (1994, 1998)]. This being the case, police officials might direct their efforts towards crimes that are most likely to result in an arrest (e.g., murders committed by family members where establishing a motive and window of opportunity are relatively easy) and away from crimes that are least likely (e.g., murders committed by strangers). It is therefore expected that $\gamma_2 < 0$. Note that this instrument is likely to be exogenous to the structural crime equation since the

[15] These weights are also provided as part of the SHR. State-specific means of the SHR instruments were used to replace any missing observations on these variables.

rate of per-capita murders in a given state will not influence the proportion of murders that are committed by strangers. Thus, this instrument is likely to explain variation in the probability of arrest as discussed above and be orthogonal to the error term in the structural murder equation.

The instruments used to identify Equation (3), the probability of conviction given arrest for murder, include the proportion of state murders committed under non-felony related circumstances and the proportion of state murders committed by non-white offenders. In general, death sentences are only imposed on offenders who have committed capital murders. Thus, we hypothesize that state courts are less likely to convict murderers whose offenses occurred under non-felonious circumstances.[16] The once-lagged value of this instrument is also included in Equation (3). It is therefore expected that δ_2, $\delta_3 < 0$.

One of the most controversial aspects of capital punishment as applied in the U.S. is its apparent over-application to minority offenders. In a review of numerous studies on the death penalty, the U.S. General Accounting Office (1990) found that individuals who murdered whites were far more likely to be sentenced to death relative to those who murdered blacks. In addition, some evidence was found to suggest that the race of the offender is often an important determinant of whether he/she receives the death penalty. Such asymmetric treatment of convicted murderers is consistent with the interest group theory of the state. The decision of state judges to over-apply (in terms of the relative severity and/or number of murders committed by a given offender) the death penalty to persons belonging to a minority group might represent the underlying discriminatory preferences of their constituency.[17]

[16] The SHR classifies its murder circumstance measure into felony-type, non-felony type, suspected felony-type, and unable to determine circumstances categories. The variable employed here is constructed from the latter three groups excluding the following specific categories: (1) Argument over money, (2) Other arguments, (3) Gangland killing, (4) Youth gang killing, (5) Institutional killing, (6) Sniper attack, (7) Suspected felony, and (8) Felony by citizen. The included categories in our measure are: (i) Abortion, (ii) Lovers triangle, (iii) Killed by babysitter, (iv) Brawl under alcohol, (v) Brawl under drugs, (vi) Hunting accident, (vii) Gun cleaning, (viii) Child playing with gun, (ix) Other negligent gun, (x) Other negligent manslaughter, (xi) Other non-felony, (xii) Felony by police, and (xiii) Unable to determine circumstance.

[17] In some cases such preferences for discrimination among the constituency might exert

For example, approximately 83% of state executions that have occurred since 1976 took place in the South, a region with a long history of problems pertaining to racial discrimination. Alternatively, a constituency's underlying preferences for discrimination maybe reflected in the use of low quality legal counsel provided by public defenders. Approximately 90% of defendants (many of whom are presumably minorities) who face the possibility of receiving a death sentence for committing murder cannot afford their own attorney, and the defense counsel provided to these persons by the state will often consist of attorneys who are inexperienced, unqualified, or not provided adequate financial resources (i.e., relative to the state prosecutor) to mount an adequate defense case [Dieter (1996)]. Of course, the net effect of inadequate defense council would be to increase the probability that a minority defendant receives a death sentence.

Any systematic discrimination in applying capital punishment toward minority offenders is controlled for by including a measure of the proportion of total state murders that are committed by non-white offenders. The once-lagged value of this instrument is also included in Equation (3). Given the above discussion it is expected that δ_4, $\delta_5 > 0$. Finally, note again that the instruments used to identify Equation (3) are likely to be exogenous to the crime equation for similar reasons as to why the instrument in Equation (2) is exogenous. That is, the rate of per-capita murders in a given state would not directly determine the distribution of murders arising between felonious and non-felonious or the distribution between white and non-white offenders.

Now consider Equation (4), the contemporaneous probability of execution given conviction. The instruments used in the identification of this deterrence probability include an indicator of whether a prisoner was released from death row in the previous year (due to doubts about his/her guilt) and an indicator of whether there was a botched execution in the previous year. With respect to the former, the hypothesis is that the release of a prisoner(s) from death row in the previous year will be negatively correlated with the current year's value of Pr(e|c), or $\varphi_2 < 0$. Gist and Hill (1981) provide empirical evidence that bureaucrats appear to be risk averse and are likely to change their behavior for

---

an even more direct influence on the probability of being sentenced to death, such as in states where juries may recommend that a convicted felon receive the death penalty.

the purpose of avoiding risk. A similar rationale may apply to the case of prisoners being released from death row. For example, a prisoner's release might to lead to diminished public confidence in the administration of the particular state's penal system and lead to fewer executions being handed down if the state's Governor is risk-averse and wishes to avoid political criticism or controversy (e.g., he/she highly values being re-elected). For this instrument to be justifiably excluded from the structural murder regression it must assumed that the previous years' releases do not induce individuals to crime.

Botched executions are those in which an egregious error in the actual carrying out of an execution resulted in prolonged administration of the punishment or unnecessary pain to the executed offender. A botched execution in the previous year might affect the number of executions carried in the current year for several reasons. First, convicts awaiting their execution might file appeals for delay under the auspice of cruel and unusual punishment arguments. These efforts might serve to delay the number of executions administered in the subsequent period given the extent of judicial and legal review that would have to be carried out. Second, a botched execution might cause a state to hold all further executions until a thorough review of the execution procedures are carried out by state corrections officials or elected representatives (similar to the number of prisoners removed from death row).[18] In either case, if state bureaucrats in charge of administering the death penalty are risk averse and value not drawing attention to themselves, it is expected that $\varphi_3 < 0$.

III. Empirical Results

In this section the OLS estimates of the effects of the three deterrence probabilities (i.e., both contemporaneous and lagged) on the rate of per-capita murder is considered first. Next, the analogous results where the three contemporaneous deterrence probabilities are instrumented are presented. State and year dummies, as well as state-specific linear trends, are included in

[18] Several of these observations are associated with particularly tragic circumstances. For instance, after the 1983 botched execution of Jimmy Lee Gray in Mississippi by asphyxiation it was discovered that the executioner was intoxicated. In the state of Indiana in 1985 William E. Vandiver was pronounced dead only after five jolts of electricity administered over seventeen minutes eventually ended his life.

estimating all models. Table 2 provides descriptive statistics for the dependent variable and the various deterrence measures.

Table 2. Descriptive Statistics and Data Sources

Variable	Sample mean	Standard deviation
Per-capita murders	6.944	3.782
Pr(a)	0.785	0.318
Pr(c|a) [C,U]	0.019	0.038
Pr(c|a) [C,A]	0.019	0.039
Pr(c|a) [L,U]	0.022	0.106
Pr(c|a) [L,A]	0.022	0.106
Pr(e|c) [C,U]	0.074	0.388
Pr(e|c) [C,A]	0.067	0.363
Pr(e|c) [L,U]	0.066	0.365
Pr(e|c) [L,A]	0.050	0.304
Per-capita prisoners	208.937	117.447
Per-capita police	254.537	49.127

Note: Per-capita variables expressed per 100,000 persons. Per-capita income in thousands of (nominal) dollars per state resident. C = contemporaneous annual value, L = once-lagged annual value, A = adjusted, U = unadjusted. All figures correspond to annual state-level data are for the years 1978-1997. Due to missing or zero-valued observations on arrests and death sentences, the actual number of observations per variable varies between 793 and 1,000. The sources of the variables used in the analysis are as follows. Murders, murder arrests, robberies, and aggregated assaults are from the U.S. Federal Bureau of Investigation, *Uniform Crime Reports* (various years). Death sentences, executions, and prisoners are from the U.S. Bureau of Justice Statistics, *Handbook of Criminal Justice Statistics* (various years). Unemployment rates, percent metropolitan population, and income are from the U.S. Census Bureau, *Statistical Abstract of the United States* (various years). Poverty rates are from the U.S. Census Bureau, available at <http://www.census.gov/hhes/poverty/histpov/hstpov21.html>. Total population, percentage age breakdowns, and percent black are from the U.S. Census Bureau, available at <http://eire.census.gov/popest/estimates.php>. Percent strangers, nonfelony, and nonwhite are from the U.S. Federal Bureau of Investigation, *Supplementary Homicide Reports* (various years). Botched executions and releases are from the Death Penalty Information Center, available at <www.deathpenaltyinfo.org>.

A. Ordinary Least Squares Estimates

Table 3 presents the OLS (structural) estimates. The first and second employ the contemporaneous deterrence probability measures, defined in unadjusted and adjusted terms, respectively. In each case the signs of the estimated coefficients on the three deterrence probabilities are negative. However, only the estimated coefficients on Pr(a) are statistically significant at conventional levels.

Table 3. Ordinary Least Squares Estimates of Per-Capita Murders

	Contemporaneous models		Lagged models	
	Unadjusted probability	Adjusted probability	Unadjusted probability	Adjusted probability
Pr(a)	-0.61***	-0.54***	-0.41**	-0.36**
	(3.94)	(3.84)	(2.54)	(2.51)
Pr(c|a)	-1.53	-1.17	-0.39	-0.11
	(1.07)	(1.17)	(0.50)	(0.17)
Pr(e|c)	-0.10	-0.13	-0.09	-0.12
	(1.24)	(1.62)	(1.05)	(1.34)
Per-capita prisoners	-0.01***	-0.01***	-0.01***	-0.01***
	(5.24)	(5.32)	(5.38)	(5.42)
Per-capita police	-0.00427	-0.00394	-0.01*	-0.01*
	(1.21)	(1.31)	(1.79)	(1.82)
% Unemployed	-0.18***	-0.13***	-0.19***	-0.14***
	(4.72)	(4.06)	(4.82)	(4.38)
Per-capita income	0.31**	0.32***	0.29**	0.30**
	(2.30)	(2.74)	(2.03)	(2.52)
% Metro	0.000786	-0.000845	-0.00207	0.00222
	(0.03)	(0.03)	(0.05)	(0.09)
% Poverty	-0.03	-0.02	-0.03	-0.01
	(1.37)	(0.99)	(1.14)	(0.52)
% Black	0.73**	0.81***	0.97***	0.89***
	(2.41)	(2.97)	(3.21)	(3.48)

Table 3. (Continued) Ordinary Least Squares Estimates of Per-Capita Murders

	Contemporaneous models		Lagged models	
	Unadjusted probability	Adjusted probability	Unadjusted probability	Adjusted probability
% Ages 18-24	-0.00363	0.08	0.06	0.12
	(0.03)	(0.75)	(0.41)	(1.08)
% Ages 25-44	0.17	0.24	0.16	0.24
	(0.96)	(1.55)	(0.90)	(1.56)
% Ages 45-64	0.79***	0.62***	0.78***	0.62***
	(2.94)	(2.70)	(2.84)	(2.70)
% Ages 65 and over	0.31	0.16	0.21	0.14
	(1.37)	(0.81)	(0.93)	(0.69)
Per-capita robberies	0.02***	0.02***	0.02***	0.02***
	(16.64)	(18.44)	(16.49)	(19.18)
Per-capita assaults	0.00298***	0.00380***	0.00323***	0.00357***
	(3.21)	(4.45)	(3.41)	(4.22)
Constant	-30.47**	-29.01***	-28.80**	-28.47***
	(2.58)	(2.91)	(2.38)	(2.89)
Observations	781	970	776	974
F	137.27***	161.85***	133.84***	165.80***
Adjusted R-squared	0.96	0.96	0.96	0.96

Note: The dependent variable is the number of state murders per 100,000 state residents. The data set is comprised of annual state level data from 1978-1997. Absolute values of t-statistics in parentheses. * denotes statistical significance at the 10% level; ** significance at the 5% level; and *** significance at the 1% level in a two-tailed test. Estimated coefficients and t-statistics on state indicators, year indicators, and state-specific time trends not shown.

The estimated coefficients on the other law enforcement measures, per-capita police and prison populations, are both negative with the former being statistically significant in both specifications. The state unemployment rate is negatively correlated with per-capita murders and is statistically significant in both specifications. This finding is consistent with Raphael and Winter-Ebmer (2001). Per-capita income is positive and statistically significant in both specifications. The percent residing in metro areas and percent in poverty are not statistically significant at conventional levels in either case. However, the percent black is positive and statistically significant in both specifications. The various signs and significance of the age structure variables are similar in both instances. Finally, in either specification the rates of per-capita robbery and aggravated assault are positive and statistically significant determinants of the per-capita rate of murder.

The third and fourth columns of Table 3 employ the lagged deterrence probabilities. Again, the former employs the unadjusted probabilities and the latter the adjusted probabilities. Once again the only statistically significant deterrence probability is Pr(a), although all the deterrence probabilities carry negative signs. The estimated coefficients on per-capita prisoners and police are again negative with the latter now statistically significant at the 10% level in both specifications. Results with respect to the remaining covariates are similar to those found in the specifications employing the contemporaneous deterrence probabilities.

B. Two Stage Least Squares Estimates

As discussed previously, causality between the per-capita rate of murder and the contemporaneous deterrence probabilities might be bi-directional. This would be the case if state authorities execute more offenders in response to increasing rates of murder or if the response of rational murders to the death penalty is to eliminate potential witnesses (i.e., for the purpose of lowering detection and evading capital punishment). As such, the estimates on the contemporaneous deterrence probabilities in Table 3 might suffer from simultaneity bias and be either over or under-estimated depending on the relative magnitudes of the deterrent versus lethality effects.

Table 4 presents the results obtained by estimating the simultaneous system consisting of Equations (1)-(4) where the contemporaneous deterrence probabilities are instrumented.[19] Instrumentation leads to a more negative

[19] All first and second-stage estimates are also obtained by weighting observations by state population.

Table 4. Two-Stage Least Squares Estimates of Per-Capita Murders

	Unadjusted probabilities	Adjusted probabilities
Pr(a)	-2.57***	-2.16**
	(2.60)	(2.08)
Pr(c|a)	-1.45	-7.43
	(0.12)	(0.77)
Pr(e|c)	-1.48**	-2.15***
	(2.24)	(2.94)
Per-capita prisoners	-0.01***	-0.01***
	(4.04)	(3.87)
Per-capita police	-0.00163	-0.000932
	(0.34)	(0.21)
% Unemployed	-0.18***	-0.11**
	(3.58)	(2.18)
Per-capita income	0.34*	0.32*
	(1.89)	(1.88)
% Metro	0.01	0.00
	(0.27)	(0.08)
% Poverty	-0.01	0.01
	(0.19)	(0.32)
% Black	1.87***	2.29***
	(3.32)	(3.78)
% Ages 18-24	-0.09	0.08
	(0.46)	(0.49)
% Ages 25-44	0.09	0.11
	(0.31)	(0.45)
% Ages 45-64	1.16***	0.90**
	(2.81)	(2.57)
% Ages 65 and over	0.40	0.13
	(1.33)	(0.44)
Per-capita robberies	0.02***	0.02***
	(11.83)	(12.11)
Per-capita assaults	0.0044***	0.01***
	(2.93)	(3.77)

Table 4. (Continued) Two-Stage Least Squares Estimates of Per-Capita Murders

	Unadjusted probabilities	Adjusted probabilities
Constant	-39.57 **	-34.24 **
	(2.20)	(2.38)
Observations	781	970
F	80.12 ***	80.56 ***
χ^2(OIR)	7.97	17.57
p(OIR)	0.16	0.00
χ^2(DWH)	5.48	10.12
p(DWH)	0.00	0.00

Note: The dependent variable is the number of state murders per 100,000 state residents. The data set is comprised of annual state level data from 1978-1997. Absolute values of t-statistics in parentheses. * denotes statistical significance at the 10% level; ** significance at the 5% level, and *** significance at the 1% level in a two-tailed test. Estimated coefficients and t-statistics on state dummies, year dummies, and state-specific time trends not reported.

estimated coefficient on Pr(a). This result is consistent with reverse causation operating in the positive direction on this measure.

The 2SLS coefficient estimates with respect to Pr(c|a) remain negative and statistically insignificant, but are substantially larger in magnitude (i.e., more negative) relative to their OLS counterparts. Thus, Pr(c|a) is found to have no influence on the rate of per-capita murders (i.e., whether or not the deterrence probabilities are defined in a contemporaneous or lagged manner). In particular, note that this result holds even when the contemporaneous measures of Pr(c|a) are instrumented.

The proxy for the subjective probability of execution given being sentenced to death, Pr(e|c), becomes more negative and turns statistically significant in both the non-corrected and corrected 2SLS specifications relative to the comparable OLS estimates. The 2SLS coefficient ranges from -1.48 in the unadjusted probabilities model to -2.15 in the adjusted probabilities case. These estimates imply that a state execution deters, on average, approximately

14 and 19 murders per year, respectively.[20] The 95% confidence interval around the mean for the non-adjusted estimate is approximately [4, 25] and [7, 31] for the adjusted estimate. Taking the unadjusted probabilities model as the preferred case,[21] these results indicate that each execution deters at least 4 murders per year and at most 25 murders. Interestingly, this range includes estimates obtained from both the original Ehrlich (1975) study (who found 18 murders being deterred per execution) and more recent studies employing panel data estimation techniques.[22]

The estimated coefficients on per-capita prison populations and police remain negative although the latter is no longer statistically significant at conventional levels. Per-capita income is again found to be positive and statistically significant. Like their OLS counterparts the coefficient estimates on the percent metropolitan population percent poverty are statistically insignificant at conventional levels. The percent black is again found to be positive and statistically significant. All remaining explanatory variables have similar signs and levels of statistical significance to the corresponding OLS estimates.

C. Tests of the Instruments

Since the number of instruments exceeds the number of endogenous regressors a joint test of the model's specification and overidentifying restrictions can be conducted.[23] The computed χ^2 statistics and the associated

---

[20] This result is calculated as $\beta_{\Pr(e/c)} * (Pop_{1997} / Sen_{1996})$ where Pop_{1997} is the total population (in thousands) of the executing states in 1997 and Sen_{1996} is the number of persons sentenced to death in 1996.

[21] Recall that the adjusted probabilities model does not satisfy the test of overidentifying restrictions. Thus, emphasis is placed on the results obtained from the unadjusted probabilities model.

[22] Using county-level panel data, Dezhbakhsh et al. (2003) estimate that an execution deters approximately 18 murders on average. Mocan and Gittings (2003), using state-level data, estimate five murders being deterred per execution.

[23] The test statistic is calculated as the product of the sample size and the uncentered R^2 from a regression of the second stage residuals on the entire set of explanatory variables and instruments. This test statistic is distributed χ^2 with degrees of freedom equal to the

p-values of the tests of overidentifying restrictions are presented in Table 4. In the unadjusted probabilities specification the p-value is within conventional bounds. This supports both the model's specification and exogeneity of the identifying instruments. On the other hand, in the adjusted probability specification the low p-value (0.00) rejects the overidentifying restrictions. Despite this rejection, both models produce point estimates of comparable sign, size, and statistical significance for all the explanatory variables. In all instances the F-tests of the joint significance of the excluded instruments reject the null hypothesis that the coefficients are zero at well above the 95% level of confidence. As such, the estimates are unlikely to exacerbate the bias resulting from reliance on a finite sample of observations.[24]

Finally, note that the influential review of Brier and Fienberg (1980, pp. 178-179) criticizes the treatment of the deterrence probabilities as endogenous variables.[25] The authors argue that it is inappropriate to model murder rates as affecting the deterrence probabilities since any reverse causation will only occur with a lag. As such, they regard the issue of identification as moot and, therefore, the use of simultaneous equations methods unnecessary. Note, however, that their criticism is only relevant to the theory that governing authorities can and will conduct more sentences or executions in response to higher murder rates. Clearly, the very nature of the judicial process governing state executions prohibits such a feedback from occurring. However, as stated previously a lethality effect of capital punishment might be operative as well. This effect might introduce endogeneity into the contemporaneous deterrence probabilities and thus validate the use of simultaneous equations methods.

---

difference between the number of identifying instruments and the number of endogenous regressors (in this instance five).

[24] These results and the first-stage regressions are available upon request.

[25] In actuality, their original criticism was directed at Ehrlich (1975) who only treated the probability of apprehension and probability of conviction given apprehension as endogenous. However, the issues raised by Brier and Feinberg (1980) also apply to the treatment of Pr(e|c) as endogenous. Indeed, an offender in the U.S. is unlikely to be executed in the same year that their offense was committed. As such, the treatment of Pr(e|c) as an endogenous variable might be questionable regardless of whether it is the deterrent effect or lethality effect that is actually operative.

In response to the concerns raised by Brier and Fienberg (1980), Durbin-Wu-Hausman (DWH) tests are performed to examine whether it is appropriate to treat the three deterrence probabilities as endogenous variables.[26] The results of the DWH endogeneity tests are also presented in Table 4. The high (low) values of the F-statistic (p-value) [5.48 (0.00) in the unadjusted probabilities model and 10.12 (0.00) in the adjusted probabilities model] indicate that the deterrence probabilities are most appropriately treated as statistically endogenous explanatory variables. Therefore, the OLS estimates of the deterrence probabilities in Table 3 are likely to suffer from simultaneity bias.[27]

IV. Discussion

The empirical results obtained from the linear specifications demonstrate substantial differences between models that employ lagged as opposed to contemporary measures of the deterrence probabilities and, in the context of the latter, those that instrument the probabilities and those that treat them as exogenous. The question now becomes whether these differences can be rationalized in practical terms. In other words, how might these rather dramatic differences (to the extent they are to be believed) be explained?

The differences between the lagged and contemporaneous linear models seems to indicate that any deterrent effect of capital punishment, to the extent that it exists, is likely to effect the rate of murder initially (i.e., within the

[26] This method involves the regression of each of the contemporaneous execution probabilities on the exogenous variables contained in Equation (1) and the entire set of excluded instruments contained in Equations (2)-(4). The predicted values obtained from these regressions are then included in the specification of Equation (1) as additional explanatory variables. A joint F-test of the significance of the included predicted values is then performed. A high value of the corresponding F-statistic (conversely a low p-value) rejects the null hypothesis of the exogeneity of the execution probabilities. This in turn would indicate that the execution probabilities are endogenous, and that simultaneous equation methods are indeed appropriate for estimation.

[27] A double-logarithmic specification of the 2SLS was also estimated (these results are available upon request). In this specification, the coefficient estimates on Pr(c|a) and Pr(e|c) remain negative but turn statistically insignificant The majority of the other coefficient estimates remain comparable to their linear 2SLS counterparts.

particular year a given execution(s) take place) but then tends to dampen quickly.[28] As such, the deterrent effect of capital punishment appears to arise from the process of administering executions and not from the existence of a death penalty law. In other words, executions appear to deter murder only through their announcement, i.e., if potential murders do actually witness an execution in proximity to the time in which they plan on committing their offense, then they will be less likely to commit a homicide. On the other hand, having a death penalty provision on the books but not meting out executions will not force potential offenders to update their subjective probability assessments and do little to deter the rate of murder.

Differences within and between the deterrence probabilities' estimates in the linear 2SLS models also support the hypothesis of an announcement effect regarding state executions. Specifically, note that while Pr(e|c) is negative and statistically significant in both cases, Pr(c|a) is never statistically different from zero. Again, one possible explanation for this discrepancy is the relative extent to which these probability measures are observable to potential murderers. Arguably, it is very difficult for a potential murderer (and probably most persons) to ascertain the number of individuals who have been or will be convicted of murder in the current year (or any other). On the other hand, information regarding policing efforts and executions are much more likely to be publicly disseminated through the media, word of mouth, and similar information channels. Thus, the results suggest that potential offenders are more likely to respond to variations in these factors than they are to changes in largely unobservable judicial action.

Several caveats to the analysis must be stated. First, as mentioned earlier, many previous studies that have attempted to estimate the deterrent effect of capital punishment have employed a double-logarithmic specification to the structural murder equation. The calculation of the deterrent effect of capital punishment from these studies is typically derived from an elasticity that takes values below one in absolute terms. The consequence of this is that executions will necessarily be subject to diminishing returns at higher rates of execution. On the other hand, the functional form chosen for the structural murder equation in this paper is linear and, again, chosen for the purpose of being

[28] This appears to apply to the probability of arrest as well, although to a lesser degree.

forced to assign ad hoc positive values of the deterrence probabilities when they take values of zero. The use of a linear model (in conjunction with the relatively narrow range of observed variation in the deterrence probabilities) might imply that the estimated lives for a life multiplier of 14 may not apply to higher frequencies of execution. As such, caution needs to be taken in extrapolating the implications of the estimates provided here.

Second, the results suggest that further attention should be given to the long-established argument that the death penalty is over-applied to minorities. Indeed, in the first-stage estimates the contemporaneous percentage of state murders committed by minority offenders was found to be positive and statistically significant determinant of the probability of being sentenced to death. Of course, this may simply reflect the tendency for minorities to commit the most serious forms of murder (e.g., killing multiple victims, offenses against juveniles, etc.). On the other hand, black offenders are far more likely to receive the death penalty when the victim is white than when the victim is black [U.S. GAO (1990)]. As such, future research employing data with greater granularity with respect to the severity of murder(s) committed and race of the offender(s) and victim(s) would be useful in addressing concerns regarding the unjustness of the death penalty's application.

Finally, the results appear to be highly sensitive to functional form. When the simultaneous equations model is specified in double-logs the estimated deterrent effect of capital punishment disappears. While other recent studies report a deterrent effect of capital punishment using either linear or logarithmic functional forms [e.g., Dezbakhsh et al. (2003), Mocan and Gittings (2003)], these estimates nonetheless highlight the longstanding difficulty in conclusively determining whether or nor capital punishment deters murder, a difficulty which is unlikely to be resolved anytime soon.

V. Conclusion

A panel of state-level data over the years 1978-1997 is employed to estimate the deterrent effect of capital punishment. Specific attention is paid to estimation in light of two forms of potential endogeneity bias: unobserved heterogeneity bias arising from omitted structural factors that determine both the rate of murders and executions concurrently, and simultaneity bias that results from the effect the rate of murder exerts on the various execution probabilities. To overcome these statistical problems, per-capita murder equations are estimated that control for sources of unobserved state and year-specific heterogeneity (through the estimation of fixed-effects models) and the effects of reverse causality (via the use of a simultaneous equations model which exploits specific identifying restrictions motivated from application of the theory of public choice to the operation of the criminal justice system and bureaucratic behavior). The estimates of the deterrent effect of state executions appear to be relatively robust to model specification. Besides controlling for the effects of a large number of unobservable factors, other time-variant determinants of crime including prison populations, police employment, economic factors, demographics, and possible substitution bias arising from murder being the by-product of other crimes are also controlled for. In addition, the employed set of instruments is found to have relatively strong explanatory power while tests of overidentifying restrictions provide general support for the models' specification.

The results also indicate that (linear) OLS estimates of the deterrence probabilities do in fact suffer from simultaneity bias and underestimate the deterrent effect of state executions. This implies that the deterrent effect arising from executing convicted offenders more than offsets any corresponding lethality effects that may result from the rational response of offenders to commit more murders. Specifically, it is estimated that each state execution deters somewhere between 4 and 25 murders per year (14 being the average). Assuming that the value of a human life is approximately $5 million [i.e., the average of the range provided by Viscusi (1993)], the estimates imply that on average each execution results in society avoiding the loss of approximately $70 million per year, all else equal (i.e., ignoring all other corresponding social benefits and costs of implementing capital punishment). Finally, the results suggest that the announcement effect of capital punishment, as opposed to the existence of a death penalty provision, is the mechanism actually driving the deterrent effect associated with state executions.

It must also be noted that even if capital punishment is a deterrent it does not follow that capital punishment should be imposed. The apparent sentencing of innocent persons to death in the U.S. marks a serious flaw with the system of capital punishment, and further measures must be implemented to ensure that such mistakes do not continue.

References

Andreoni, J. (1995), "Criminal Deterrence in the Reduced Form: A New Perspective on Ehrlich's Seminal Study," *Economic Inquiry* **33**: 476-483.

Becker, G.S. (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy* **76**: 169-217.

Bedau, H. (1997), *The Death Penalty in America: Current Controversies*, Boston, MA, Oxford University Press.

Blumstein, A., Cohen, J., and D. Nagin (1978), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates, Report of the Panel of Deterrence and Incapacitation*, Washington, DC, National Academy Press.

Benson, B.L. (1995), "Understanding Bureaucratic Behavior: Implications from the Public Choice Literature," *Journal of Public Finance and Public Choice* **2-3**: 89-117.

Benson, B.L. (1999), "An Economic Theory of the Evolution of Governance and the Emergence of the State," *Review of Austrian Economics* **12**: 131-160.

Benson, B.L., Kim, I., and D.W. Rasmussen (1994), "Estimating Deterrence Effects: A Public Choice Perspective on the Economics of Crime Literature," *Southern Economic Journal* **61**: 161-167.

Benson, B.L., Rasmussen, D.W., and I. Kim (1998), "Deterrence and Public Policy: Trade-Offs in the Allocation of Police Resources," *International Review of Law and Economics* **18**: 77-100.

Bound, J., Jaeger, D.A., and R.M. Baker (1995), "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association* **90**: 443-450.

Brier, S.S., and S.E. Feinberg (1980), "Recent Econometric Modelling of Crime and Punishment: Support for the Deterrence Hypothesis?," *Evaluation Review* **4**: 147-191.

Brumm, H.J., and D.O. Cloninger (1996), "Perceived Risk of Punishment and the Commission of Homicides: A Covariance Structure Analysis," *Journal of Economic Behavior and Organization* **31**: 1-11.

Cameron, S. (1994), "A Review of the Econometric Evidence on the Effects of Capital Punishment," *Journal of Socio-Economics* **23**: 197-214.

Dezhbakhsh, H., Rubin, P.H., and J.M. Shepherd (2003), "Does Capital Punishment Have a Deterrent Effect? New Evidence from Post-moratorium Panel Data," *American Law and Economics Review* **5**: 344-376.

Dieter, R.C. (1996), "Twenty Years of Capital Punishment: A Re-evaluation," Death Penalty Information Center, mimeo.

Donohue, J.J., (2003), "Divining the Impact of Concealed Carry Laws," in J. Ludwig and P. Cook, eds., *Evaluating Gun Policy: Effects on Crime and Violence,* Washington, DC, Brookings Press.

Duggan, M. (2001), "More Guns, More Crime," *Journal of Political Economy* **109**: 1086-1114.

Ehrlich, I. (1975), "The Deterrent Effect of Capital Punishment: A Question of Life and Death," *American Economic Review* **65**: 468-474.

Ehrlich, I. (1977), "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence," *Journal of Political Economy* **85**: 741-788.

Gist, J.R., and R.C. Hill (1981), "The Economics of Choice in the Allocation of Federal Grants: An Empirical Test," *Public Choice* **36**: 63-73.

Grogger, J. (1990), "The Deterrent Effect of Capital Punishment: An Analysis of Daily Homicide Counts," *Journal of the American Statistical Association* **85**: 295-303.

Katz, L., Levitt, S.D., and E. Shustorovich (2003), "Prison Conditions, Capital Punishment, and Deterrence," *American Law and Economics Review* **5**: 318-343.

Lott, J.R., and D.B. Mustard (1997), "Crime, Deterrence, and Right-to-Carry Concealed Handguns," *Journal of Legal Studies* **26**: 1-68.

Lott, J.R., and J. Whitley (2003), "Measurement Error in County-Level Data," *Journal of Quantitative Criminology* **19**: 185-198.

Maltz, M.D., and J. Targonski (2002), "A Note on the Use of County-Level UCR Data," *Journal of Quantitative Criminology* **18**: 297-318.

Maltz, M.D., and J. Targonski (2003), "Measurement and other Errors in

County-Level UCR Data: A Reply to Lott and Whitley," *Journal of Quantitative Criminology* **19**: 199-206.

Marvell, T.B., and C.E. Moody (2001), "The Lethal Effects of Three-Strikes Laws," *Journal of Legal Studies* **30**: 89-106.

Mocan, H.N., and R.K. Gittings (2003), "Getting Off Death Row: Commuted Sentences and the Deterrent Effect of Capital Punishment," *Journal of Law and Economics* **46**: 453-478.

Plassmann, F., and T.N. Tideman (2001), "Does the Right to Carry Concealed Handguns Deter Countable Crimes? Only a Count Analysis Can Say," *Journal of Law and Economics* **44**: 771-798.

Raphael, S., and R. Winter-Ebmer (2001), "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics* **44**: 259-284

Sah, R.K. (1991), "Social Osmosis and Patterns of Crime," *Journal of Political Economy* **99**: 1272-1295.

U.S. General Accounting Office (1990), "Death Penalty Sentencing: Research Indicates Pattern of Racial Disparities," report GAO-GDD-90-57.

Viscusi, W.K. (1993), "The Value of Risks to Life and Health," *Journal of Economic Literature* **31**: 1912-1946.